Electrical Engineering

Safe and Robust Reinforcement Learning for Autonomous Cyber-Physical Systems

Nathaniel Hamilton

Dissertation under the direction of Professor Taylor T. Johnson

Recent successes have brought Reinforcement Learning (RL) to the forefront of the AI discussion. Additional research consistently shows RL training produces optimal results even when trained with inaccurate or incomplete models. Instead of hard-coding, agents are programmed via reward and punishment without needing to specify how the task is to be achieved. Because of this, RL techniques are extremely attractive for autonomous cyber-physical systems applications, which often have complex dynamics and environments that are difficult to model. However, the behaviors of these RL systems are often difficult to interpret and predict. This is unacceptable for safety-critical systems, where unpredictable and unsafe behavior could be the difference between life and death.

To protect systems, and speed up the training process, engineers can train their RL agents in simulation and then transfer the learned control policy to a real-world system. This process is a challenging problem referred to as the sim2real challenge. Simulation environments abstract away a lot of the nuance and noise of the real world. As a result, RL agents trained in simulation can end up "brittle," i.e., when confronted with scenarios that differ from the examples seen in training, they fail to contextualize the situation and break. Thus, sim2real transfers often have catastrophic results where the real-world results are drastically different from simulation. Therefore, if we are to take advantage of the many benefits of training in simulation, we must emphasize training safer and more robust agents and/or formally

verifying their behavior before deploying on safety-critical, real-world systems.

To this end, the following dissertation presents research on methods for making safer and more robust RL agents tasked with controlling autonomous cyber-physical systems. First, we demonstrate some advantages RL has over other machine learning approaches for addressing the sim2real challenge. Next, we explore various approaches for ensuring agents are trained to ensure safe behavior during and after training. Finally, we present our tool for incorporating safety in reward generation by converting written specifications into reward functions.